# Building disease progression models from longitudinal biomarkers

Clémentine Fourrier, Igor Koval, Simona Bottani, Olivier Colliot, Stanley M Durrleman

▶ **To cite this version:**

Clémentine Fourrier, Igor Koval, Simona Bottani, Olivier Colliot, Stanley M Durrleman. Building disease progression models from longitudinal biomarkers. 5th Annual Human Brain Project Summit, Oct 2017, Glasgow, United Kingdom. hal-02466386

**HAL Id: hal-02466386**
**https://hal.science/hal-02466386**

Submitted on 4 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BUILDING DISEASE PROGRESSION MODELS FROM LONGITUDINAL BIOMARKERS

Human Brain Project
Unifying our understanding of the human brain.

European Commission

Brain diseases affect one third of the European population. It is crucial to improve our understanding of those diseases to better manage treatments and alleviate patients' symptoms, as well as, on a longer timescale, better understand the human brain on its whole.

For this purpose, SP8 is building the *Medical Informatics Platform* (MIP). It will provide an infrastructure to connect hospital databases and clinicians to researchers and their algorithms (while respecting patient confidentiality).
At the end of SGA1, this platform will already contain several research algorithms which will contribute to characterizing the disease pathways and progressions at different scales, among which the one introduced in this poster.

Clementine FOURRIER, Igor KOVAL, Simona BOTTANI, Olivier COLLIOT, Stanley DURRLEMAN – ARAMIS, Brain and Spine Institute, Paris

## OBJECTIVES

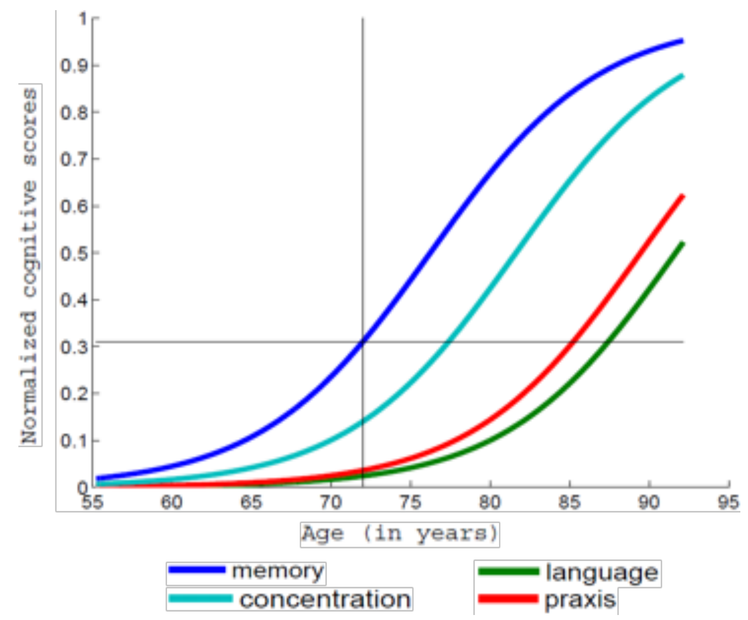
This task implements an algorithm, whose aim is to **build disease progression models from longitudinal measurements** (such as clinical data, cognitive scores, biological or neuroimaging markers).

From a population of patients, the model learns :
- a **group-average trajectory**, characterizing the ordering and dynamics of the alterations
- **individual trajectories for each patient**, modeling the variability of the dynamic between patients

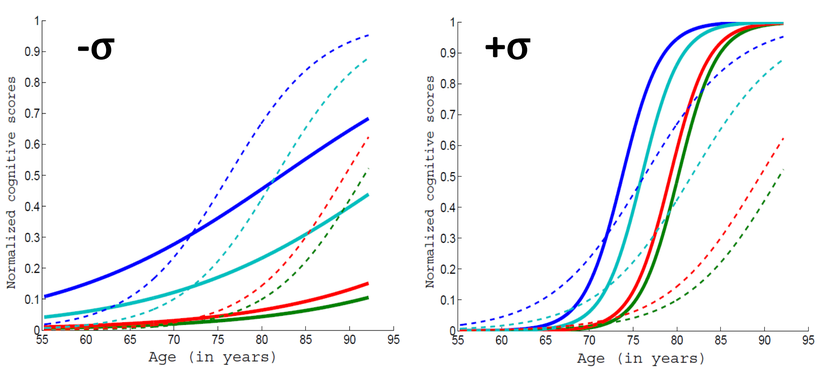
This shall allow users to better understand the disease progression, but also predict the evolution of patients and to adjust treatments accordingly.

## EXAMPLE OF CLINICAL APPLICATION

**Average trajectories for a set of cognitive scores for patients with Alzheimer Disease (AD)**, from a normal state to an abnormal state. From the data of 248 patients from ADNI, with on average 6 temporal points per patient. [1]




Example of the **population variability** for one of the computed parameters (the acceleration factor, linked to the pace and age at onset). *In dotted, the average, in bold, the variability.*




The model also **accounts for temporal variability** (disease pace and age at onset) **and spatial variability** (ordering) within the population

## METHODS

This component relies on a highly innovative statistical learning approach [1, 2].

Its core is a **new non-linear mixed effects model for Riemannian data**.

$$\mathbf{y}_{i,j} = \eta^{\mathbf{w}_i}(\boldsymbol{\gamma}, \psi_i(t_{i,j})) + \boldsymbol{\varepsilon}_{i,j}$$

- $y_{i,j}$ the output data for patient i, at observation j
- $t_{i,j}$ the value of the data for patient i, at observation j

} Data

- $\boldsymbol{\gamma}$ the average trajectory
- $\psi_i$ the time reparametrization
- $\boldsymbol{\eta}^{wi}$ the spatial variability (parallel to the average trajectory)

} Computed parameters

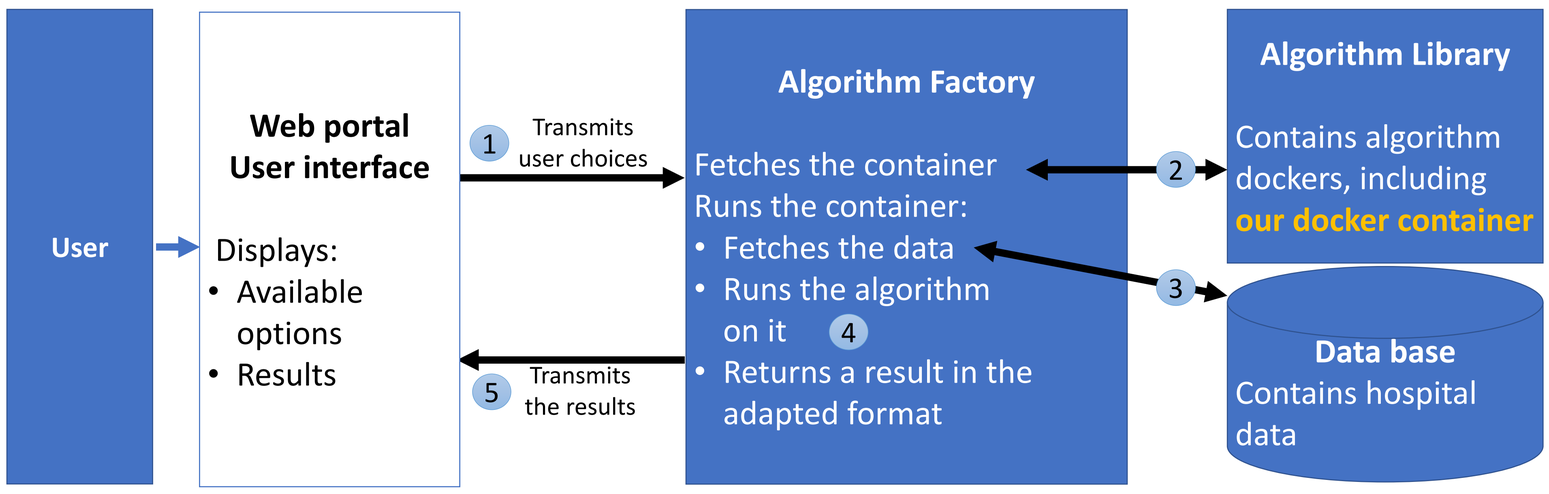
The model learns a typical group-average trajectory. The subject-specific trajectories are defined via spatial and temporal transformations of the group-average scenario.

This model does not require an absolute reference time for the disease trajectories. This is critical for neurological diseases because, at the same age, patients can/will be at different stages.
Finally, the **parameters estimation** is based on the **MCMC-SAEM algorithm** (or Monte Carlo Markov Chain - Stochastic Approximation Expectation Maximization algorithm).

## IMPLEMENTATION & INTEGRATION

- **Implementation**:
  - C++ (chosen for computational efficiency, with a modular architecture to facilitate code expansion)
  - Libraries: Linear Algebra with Armadillo, Unit Testing with GoogleTest
  - Visualisation: Python with matplotlib and numpy
- **Integration in the MIP**: Python Docker container (works as a standalone package, configured to contain everything needed to run a software: dependencies, environment, …)

Docker containers are a very portable and secure way to share executables, since you provide a pre-configured standalone « box » which can be directly run without any set up. Here, the container provided by the MIP teams can be connected to the databases when run, and was extended to include our algorithm and its needed dependencies.




## CONCLUSION

The ICM is developing a software component for SP8, based on an **innovative and powerful algorithm**, to allow clinicians to **test their hypotheses on disease progression, and predict patient evolution**.

In SGA2, this algorithm will include personnalisation, and networks propagation models.

[1] Jean-Baptiste Schiratti, Stéphanie Allassonniere, Olivier Colliot, Stanley Durrleman. Learning spatiotemporal trajectories from manifold-valued longitudinal data. Neural Information Processing Systems, Dec 2015, Montréal, Canada. Advances in Neural Information Processing Systems 28
[2] Jean-Baptiste Schiratti, Stéphanie Allassonniere, Olivier Colliot, Stanley Durrleman. A Bayesian mixed-effects model to learn trajectories of changes from repeated manifold-valued observations. The Journal of Machine Learning Research, In Press

TASK OWNER

ARAMIS LAB BRAIN DATA SCIENCE

ICM Institut du Cerveau et de la Moelle épinière

INSTITUTIONS

cea

מכון ויצמן למדע WEIZMANN INSTITUTE OF SCIENCE

Ernst Strüngmann Institute In Cooperation with Max Planck Society

UMEÅ UNIVERSITET

אוניברסיטת חיפה University of Haifa

Champalimaud Foundation

SPONSORS

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

UCL